CLAIMS

1.   A method providing computer-implemented content propagation for enhanced document retrieval, the method comprising:

identifying reference information directed to one or more documents, the reference information being identified from one or more sources of data independent of a data source comprising the one or more documents;

extracting metadata that is proximally located to the reference information;

calculating relevance between respective features of the metadata to content of associated ones of the one or more documents;

for each document of the one or more documents, indexing associated portions of the metadata with the relevance of features from the respective portions into original content of the document; and

wherein the indexing generates one or more enhanced documents.

2.   A method as recited in claim 1, wherein the reference information comprises a link and/or substantially unique document ID associated with a document of the one or more documents.

3.   A method as recited in claim 1, wherein the one or more documents are knowledge base article(s), product help, task, and/or developer data.

4.   A method as recited in claim 1, wherein the one or more sources of data comprise service request(s), newsgroup posting(s), and/or search query log(s).

5.   A method as recited in claim 1, wherein the metadata is semantically and/or contextually related to associated ones of the one or more documents.

6.   A method as recited in claim 1, wherein the metadata comprises a title of a document, product problem context, and/or product problem resolution information.

7.   A method as recited in claim 1, wherein for each enhanced document of the one or more enhanced documents, there is a corresponding original document from which the enhanced document was generated.

8.   A method as recited in claim 1, wherein calculating the relevance is based on how many times a particular document of the one or more documents is identified within its context in the metadata.

9.   A method as recited in claim 1, wherein the metadata comprises article title(s), product problem context, and/or product problem resolution information, and wherein calculating relevance further comprises weighting the article title(s) and/or product problem context to indicate a greater relevance than any product problem resolution information.

10. A method as recited in claim 1, wherein calculating relevance further comprises assigning greater relevance to feature(s) of the metadata that occur in content of the data source with greater frequency as compared to the frequency of occurrence of other metadata features in the content.

11. A method as recited in claim 1, wherein calculating relevance further comprises assigning greater weight to feature(s) of the metadata found in a document of the one or more documents as a function of an age of the document.

12. A method as recited in claim 1, wherein the one or more sources of data comprise a search query log, and wherein calculating relevance further comprises:

identifying search queries from the search query log, wherein the search queries have a relatively high frequency of occurrence (FOO) to search the data source;

determining article(s) selected by an end-user from search query results, the article(s) being from the data source; and

determining missing end-user selection(s), where a missing end-user selection is an article in the search query results that was not selected.

13. A method as recited in claim 12, wherein determining missing end-user selection(s) further comprises clustering heterogeneous objects using inter-layer links to determine importance measurements for features of the heterogeneous objects, the heterogeneous object comprising a first cluster of similar queries and a second cluster of related documents, the similar queries having been identified in the search query log, the similar queries being associated search result(s) comprising the one or more documents, the related documents being identified in the search result(s) independent of whether individual ones of the related documents were selected by an end-user from the search results.

14. A method as recited in claim 13, wherein the features are represented with respective nodes in the first and second clusters, and wherein the importance measurement(s) for each of the nodes is based on a similarity function that measures a distance between objects in the first and second clusters.

15. A computer-readable medium comprising computer-executable instructions providing content propagation for enhanced document retrieval, the computer-executable instructions comprising instructions for:

identifying reference information directed to one or more documents, the reference information being identified from one or more sources of data independent of a data source comprising the one or more documents;

extracting metadata that is proximally located to the reference information;

calculating relevance between respective features of the metadata to content of associated ones of the one or more documents;

for each document of the one or more documents, indexing associated portions of the metadata with the relevance of features from the respective portions into original content of the document; and

wherein the indexing generates one or more enhanced documents.

16.   The computer-readable medium of claim 15, wherein the reference information comprises a link and/or substantially unique document ID associated with a document of the one or more documents.

17.   The computer-readable medium of claim 15, wherein the one or more documents are knowledge base article(s), product help, task, and/or developer data.

18.   The computer-readable medium of claim 15, wherein the one or more sources of data comprise service request(s), newsgroup posting(s), and/or search query log(s).

19.   The computer-readable medium of claim 15, wherein the metadata is semantically and/or contextually related to associated ones of the one or more documents.

20.   The computer-readable medium of claim 15, wherein the metadata comprises a title of a document, product problem context, and/or product problem resolution information.

21.   The computer-readable medium of claim 15, wherein for each enhanced document of the one or more enhanced documents, there is a corresponding original document from which the enhanced document was generated.

22.   The computer-readable medium of claim 15, wherein calculating the relevance is based on how many times a particular document of the one or more documents is identified within its context in the metadata.

23.   The computer-readable medium of claim 15, wherein the metadata comprises article title(s), product problem context, and/or product problem resolution information, and wherein the instructions for calculating relevance further comprise instructions for weighting the article title(s) and/or product problem context to indicate a greater relevance than any product problem resolution information.

24.   The computer-readable medium of claim 15, wherein the instructions for calculating relevance further comprise instructions for assigning greater relevance to feature(s) of the metadata that occur in content of the data source with greater frequency as compared to the frequency of occurrence of other metadata features in the content.

25. The computer-readable medium of claim 15, wherein the instructions for calculating relevance further comprise instructions for assigning greater weight to feature(s) of the metadata found in a document of the one or more documents as a function of an age of the document.

26. The computer-readable medium of claim 15, wherein the one or more sources of data comprise a search query log, and wherein the instructions for calculating relevance further comprise instructions for:

identifying search queries from the search query log, wherein the search queries have a relatively high frequency of occurrence (FOO) to search the data source;

determining article(s) selected by an end-user from search query results, the article(s) being from the data source; and

determining missing end-user selection(s), where a missing end-user selection is an article in the search query results that was not selected.

27.   The computer-readable medium of claim 26, wherein the instructions for determining missing end-user selection(s) further comprise instructions for clustering heterogeneous objects using inter-layer links to determine importance measurements for features of the heterogeneous objects, the heterogeneous object comprising a first cluster of similar queries and a second cluster of related documents, the similar queries having been identified in the search query log, the similar queries being associated search result(s) comprising the one or more documents, the related documents being identified in the search result(s) independent of whether individual ones of the related documents were selected by an end-user from the search results.

28.   The computer-readable medium of claim 27, wherein the features are represented with respective nodes in the first and second clusters, and wherein the importance measurement(s) for each of the nodes is based on a similarity function that measures a distance between objects in the first and second clusters.

29. A computing device providing content propagation for enhanced document retrieval, the computing device comprising:

a processor; and

a memory coupled to the processor, the memory comprising computer-program instructions executable by the processor for:

identifying reference information directed to one or more documents, the reference information being identified from one or more sources of data independent of a data source comprising the one or more documents;

extracting metadata that is proximally located to the reference information;

calculating relevance between respective features of the metadata to content of associated ones of the one or more documents;

for each document of the one or more documents, indexing associated portions of the metadata with the relevance of features from the respective portions into original content of the document; and

wherein the indexing generates one or more enhanced documents.

30. The computing device of claim 29, wherein the reference information comprises a link and/or substantially unique document ID associated with a document of the one or more documents.

31. The computing device of claim 29, wherein the one or more documents are knowledge base article(s), product help, task, and/or developer data.

32. The computing device of claim 29, wherein the one or more sources of data comprise service request(s), newsgroup posting(s), and/or search query log(s).

33. The computing device of claim 29, wherein the metadata is semantically and/or contextually related to associated ones of the one or more documents.

34. The computing device of claim 29, wherein the metadata comprises a title of a document, product problem context, and/or product problem resolution information.

35. The computing device of claim 29, wherein for each enhanced document of the one or more enhanced documents, there is a corresponding original document from which the enhanced document was generated.

36. The computing device of claim 29, wherein calculating the relevance is based on how many times a particular document of the one or more documents is identified within its context in the metadata.

37. The computing device of claim 29, wherein the metadata comprises article title(s), product problem context, and/or product problem resolution information, and wherein the instructions for calculating relevance further comprise instructions for weighting the article title(s) and/or product problem context to indicate a greater relevance than any product problem resolution information.

38.   The computing device of claim 29, wherein the instructions for calculating relevance further comprise instructions for assigning greater relevance to feature(s) of the metadata that occur in content of the data source with greater frequency as compared to the frequency of occurrence of other metadata features in the content.

39.   The computing device of claim 29, wherein the instructions for calculating relevance further comprise instructions for assigning greater weight to feature(s) of the metadata found in a document of the one or more documents as a function of an age of the document.

40.   The computing device of claim 29, wherein the one or more sources of data comprise a search query log, and wherein the instructions for calculating relevance further comprise instructions for:

identifying search queries from the search query log, wherein the search queries have a relatively high frequency of occurrence (FOO) to search the data source;

determining article(s) selected by an end-user from search query results, the article(s) being from the data source; and

determining missing end-user selection(s), where a missing end-user selection is an article in the search query results that was not selected.

41. The computing device of claim 40, wherein the instructions for determining missing end-user selection(s) further comprise instructions for clustering heterogeneous objects using inter-layer links to determine importance measurements for features of the heterogeneous objects, the heterogeneous object comprising a first cluster of similar queries and a second cluster of related documents, the similar queries having been identified in the search query log, the similar queries being associated search result(s) comprising the one or more documents, the related documents being identified in the search result(s) independent of whether individual ones of the related documents were selected by an end-user from the search results.

42. The computing device of claim 41, wherein the features are represented with respective nodes in the first and second clusters, and wherein the importance measurement(s) for each of the nodes is based on a similarity function that measures a distance between objects in the first and second clusters.

43.   A computing device providing content propagation for enhanced document retrieval, the computing device comprising:

identifying means to identify reference information directed to one or more documents, the reference information being identified from one or more sources of data independent of a data source comprising the one or more documents;

extracting means to extract metadata that is proximally located to the reference information;

calculating means to calculate relevance between respective features of the metadata to content of associated ones of the one or more documents;

for each document of the one or more documents, indexing means to index associated portions of the metadata with the relevance of features from the respective portions into original content of the document; and

wherein the indexing generates one or more enhanced documents.

44.   The computing device of claim 43, wherein the reference information comprises a link and/or substantially unique document ID associated with a document of the one or more documents.

45.   The computing device of claim 43, wherein the one or more documents are knowledge base article(s), product help, task, and/or developer data.

46.   The computing device of claim 43, wherein the one or more sources of data comprise service request(s), newsgroup posting(s), and/or search query log(s).

47.   The computing device of claim 43, wherein the metadata is semantically and/or contextually related to associated ones of the one or more documents.

48.   The computing device of claim 43, wherein the metadata comprises article title(s), product problem context, and/or product problem resolution information, and wherein the calculating means to calculate relevance further comprise weighting means to weight the article title(s) and/or product problem context to indicate a greater relevance than any product problem resolution information.

49.   The computing device of claim 43, wherein the calculating means to calculate relevance further comprise assigning means to assign greater relevance to feature(s) of the metadata that occur in content of the data source with greater frequency as compared to the frequency of occurrence of other metadata features in the content.

50.   The computing device of claim 43, wherein the calculating means to calculate relevance further comprise assigning means to assign greater weight to feature(s) of the metadata found in a document of the one or more documents as a function of an age of the document.

51. The computing device of claim 43, wherein the one or more sources of data comprise a search query log, and wherein the calculating means to calculate relevance further comprise:

identifying means to identify search queries from the search query log, wherein the search queries have a relatively high frequency of occurrence (FOO) to search the data source;

determining means to determine article(s) selected by an end-user from search query results, the article(s) being from the data source; and

calculating means to calculate missing end-user selection(s), where a missing end-user selection is an article in the search query results that was not selected.

52. The computing device of claim 52, wherein the calculating means further comprise clustering means to cluster heterogeneous objects using inter-layer links to determine importance measurements for features of the heterogeneous objects, the heterogeneous object comprising a first cluster of similar queries and a second cluster of related documents, the similar queries having been identified in the search query log, the similar queries being associated search result(s) comprising the one or more documents, the related documents being identified in the search result(s) independent of whether individual ones of the related documents were selected by an end-user from the search results.